US Department of Education

Federal Student Aid

Integration Leadership Support Contractor

June 1, 2007

Draft Enterprise Data Management Data Policies

Final

Executive Summary

This document defines data policies necessary to promote Federal Student Aid standards and apply best practices to the creation, maintenance and sharing of business data. Adherence to these policies by all business and information technology organizations is critical to the success of Federal Student Aid efforts to manage data as an enterprise asset, and to enable more efficient business processing and improving decision-making.

The policies in this document apply to all business capability areas, Target State Vision development initiatives and operation and maintenance efforts throughout Federal Student Aid. These policies define mandatory actions by development teams and their business sponsors in the following areas:

- Data management in system development lifecycle
- Data quality program
- Data governance and enterprise metadata repository
- Use of the XML Registry and Repository
- Use of the Enterprise Conceptual Data Model and Enterprise Logical Data Model
- Master data management
- Data security and privacy

Federal Student Aid will review these policies annually, or more often if deemed necessary, for the policies to reflect the latest Federal, Department of Education and Federal Student Aid guidance.

Table of Contents

Table of Tables

Section 1: Introduction

Federal Student Aid is committed to managing data as an enterprise asset, and therefore established the Enterprise Data Management (EDM) Team. This document captures policies, defined and maintained by the EDM Team, for promoting corporate standards and applying best practices to creating, maintaining and sharing mission critical business data. Federal Student Aid will review these policies annually, or more often if deemed necessary, for the policies to reflect the latest Federal, Department of Education and Federal Student Aid guidance.

1.1 Scope and Applicability

The policies in this document cover all aspects of data management required to maximize data sharing across the enterprise and to continuously improve data quality. The policies apply to all business capability areas (BCAs), Target State Vision (TSV) development initiatives and operation and maintenance efforts throughout Federal Student Aid. The policies also apply to data transfers from partners to Federal Student Aid, such as Federal Family Education Loans (FFEL) data reporting. The EDM Team must review and validate any deviations from these policies.

1.2 Document Structure

This remainder of this document is organized into the following sections.

- **Section 2: Laws and Regulations,** lists the laws and regulations that give the Federal Student Aid EDM Team authority to set these policies.
- **Section 3: Roles and Responsibilities,** defines the stakeholder groups that are referenced in subsequent policy statements.
- **Section 4: Data Policies,** contains specific policy statements.
- **Appendix A: Acronyms**, defines all acronyms used in the document.

Section 2: Relevant Laws and Regulations

The data policies described in this document support Federal Student Aid's implementation of a series of laws and regulations. The most important of these laws and regulations include:

1. Clinger-Cohen Act (CCA) of 1996: Also known as "Information Technology Management Reform Act," the CCA provides that the government information technology (IT) departments be operated as an efficient and profitable business would be operated. Acquisition, planning and management of technology must be treated as capital investments.
2. Office of Management and Budget (OMB) Circular A-130, *Management of Federal Information Resources*, November 30, 2002: This Circular establishes policy for planning, collecting, managing, disseminating and protecting Federal information resources.
3. OMB, *Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility and Integrity of Information Disseminated by Government Agencies*, February 22, 2002, 67 FR 8452-8460: Issued pursuant to the Treasury and General Government Appropriations Act for Fiscal Year 2001 (Public Law 106-554, Section 515), the Guidelines require that agencies meet basic information quality standards. More specifically, the OMB Guidelines state that "agencies shall have a basic standard of quality (including objectivity, utility, and integrity) as a performance goal, and incorporate the standard into the agency's operations." The Guidelines require agencies to develop a process for reviewing the quality of information before information is disseminated and to report to OMB information quality complaints.
4. OMB, *Federal Enterprise Architecture* (FEA): The FEA is an OMB initiative that aims to comply with the Clinger-Cohen Act and provide a common methodology for IT acquisition in the United States federal government. It is designed to ease sharing of information and resources across federal agencies, reduce costs, and improve citizen services. The FEA is currently a collection of reference models that develop a common methodology for describing IT resources. The Data Reference Model (DRM) is one of these models, and it describes, at an aggregate level, the data and information that support government program and business line operations. This model enables agencies to describe the types of interaction and exchanges that occur between the federal government and citizens.
5. Department of Education (ED), *US Department of Education Information Quality Guidelines*"[1]: Consistent with the guidance from OMB, the Department's Information Quality Guidelines reflect the Department's policy and procedures for reviewing and substantiating the quality of information it disseminates (e.g., reports, studies, summaries), as well as for providing an administrative mechanism allowing affected persons to seek and obtain, where appropriate, correction of information not complying with the Guidelines. These Guidelines, along with those issued by OMB, represent a

[1] US Department of Education Information Quality Guidelines, http://connected.ed.gov/doc_img/infoqualguide.pdf.

performance goal for the Department and are intended only to improve the internal management of the Department.

6. *Government Performance and Results Act of 1993*, Public Law, 103-62: The purpose of the Act is to improve Federal program effectiveness and public accountability by promoting a new focus on results, service quality and customer satisfaction.

7. *Federal Information Security Management Act of 2002 (FISMA)*, 44 U.S.C. § 3541: The Act mandates Federal agencies and their affiliated parties, such as government contractors, to follow a set of mandatory processes to conduct yearly audits of computer systems. The goal of the audits is to bolster security of networks, systems, and information within the Federal government and their affiliated parties.

Section 3: Roles and Responsibilities

Data management is the responsibility of Federal Student Aid BCAs, application development teams and IT management organizations, such as the EDM Team. Following is a list of the key groups who participate in data policy implementation. These groups are referenced throughout the rest of the document. The groups defined below are general roles, not specific positions. There must be personnel assigned to perform these roles.

- Business Owners – The executive(s), program managers and/or key contacts sponsoring a given TSV initiative or an existing system, such as National Student Loan Data System (NSLDS). These individuals represent the relevant business communities and IT teams in developing and applying data policies.

- Project Managers – The managing persons with overall responsibility for delivering a given system. These individuals ensure implementation of processes and solutions in compliance with Federal Student Aid data policies. In addition, project managers evaluate and raise data issues that arise during the discovery, modeling and implementation stages of a project lifecycle.

- Data Stewards – The designated personnel from a given BCA, who coordinate data management activities to address and resolve data standardization issues and to implement data policies related to the data and/or information integrity, security, delivery and access within the assigned BCA.

- Data Architects – The data design staff from a development team who are primarily responsible for applying data policies in data discovery, data analysis and data architecture design. Data architects are also responsible for working closely with subject matter experts to uncover, document, model and validate logical information requirements into specifications that can be used to develop physical data structures.

- Database Administrators (DBA) – The individuals responsible for developing and maintaining physical database structures in accordance with related data policies. The DBAs work closely with the data architects in the creation, review, validation and expansion of associated logical and physical data models. The DBA also manages permissions based on data security standards and policies.

- The EDM Team – The centralized data management staff at Federal Student Aid with the overall responsibility of promoting best practices in managing data as an enterprise asset. The EDM Team sets standards, policies and procedures, and offers data management services to the business owners and their development projects. These services include defining best practices, sharing common metadata and providing standardized data design artifacts. The following is a list of these services:

- The EDM Team will define data management processes and procedures through which business owners, project managers, data architects and DBAs can request EDM services and can communicate data-related issues.

- The EDM Team will maintain enterprise level work products as specified in the Information work stream of the Integration Framework.[2] These work products include business use case vs. data entity CRUD matrix, operation activity information exchange diagram, and business use cases vs. information exchange matrix.

- The EDM Team will create, maintain and publish artifacts and standards in the areas of data management planning, eXtensible Markup Language (XML) elements, data analysis, data modeling, data migration, Commercial Off-the-Shelf (COTS) integration, data architecture, data integration, data quality and data security. These artifacts represent the best practices and design standards that each development project will follow.

- The EDM Team will implement and maintain the tools for data profiling, data cleansing and data quality reporting, and will provide technical know-how on using the tools. These tools will be offered as a shared technical infrastructure across the enterprise, through which the business owners can obtain a "second opinion" regarding data quality in their systems.

- The EDM Team will define and keep current data quality management best practices. The EDM Team will advise business owners on data quality requirements and expectations regarding continuous improvement of existing data quality management practices.

- The EDM Team will maintain and publish the Enterprise Conceptual Data Model (ECDM) and Enterprise Logical Data Model (ELDM). The EDM Team will provide them to business owners, project managers, data architects and DBAs as services to support data analysis, requirement gathering and physical data model development.

- The EDM Team will maintain the enterprise metadata repository (MDR), as the system of record and central access point for the enterprise data dictionary, XML core components and schemas, data sources, data movements, transformations and usage by business processes, systems, operational reports and business intelligence (BI) analyses.

- The EDM-maintained enterprise MDR will also provide capabilities to support data lineage reporting, impact analysis and service lookup.

- The EDM Team will define and publish processes and procedures for MDR users to request for retrieval and updates of metadata content.

[2] Integration Framework, Version 2, February 2007.

Section 4: Data Policies

This section contains data policies. These policies cover the following areas:

- Data management in the system development lifecycle
- Data quality program
- Data governance
- Use of XML Registry and Repository (R&R) for the Education Community
- Use of the ECDM and ELDM
- Enterprise MDR
- Master data management (MDM)
- Data security and privacy

Each of the following sub-sections starts with an introduction for a particular area and then lists the data policy statements. These policies are essential for establishing enterprise-wide data management best practices and for promoting consistent data definitions, design and architecture.

4.1 Data Management in the System Development Lifecycle

Business owners, project managers and their development team should collaborate with the EDM Team in all phases of the system development life cycle (SDLC). Such collaboration is critical for implementing standardized data definitions in system design and for setting a foundation for an integrated business and system architecture. The following policies establish such collaboration throughout the SDLC:

- **Engage the EDM Team from the initiative vision phase**: Business owners will ensure that the EDM Team is involved in the system development lifecycle from the very beginning, which is the initiative vision phase, so that the project vision reflects enterprise wide data integration considerations. The EDM Team captures these considerations by maintaining enterprise level work products as specified in the Information work stream in the Integration Framework. [3] Refer to Section 3 for a list of EDM services.

- **Conform to enterprise wide data design standards**: Project managers and data architects will refer to the enterprise wide standard data design artifacts to ensure compliance with object definitions, naming conventions, data modeling guidelines, standard data elements, XML standards and data integration interfaces. The EDM Team will maintain and provide these standard artifacts. Refer to Section 3 for a list of EDM services.

[3] Integration Framework, Version 2.0, February 2007.

- **Provide data management documentation to the EDM Team for review and approval**: Project managers and data architects will provide the EDM Team the following documentation, if applicable, for review and approval during each lifecycle phase. Note that for many of these required documents, the EDM Team will provide standard enterprise level artifacts directly to the projects.
  - Definition phase:
    - Project data management plan
    - Data migration plan
  - Construction phase:
    - Data dictionary
    - Logical data model
    - XML schemas
  - Implementation phase:
    - Physical data models
    - Data integration services design

4.2 Data Quality Program

Federal Student Aid is committed to raising data quality for both operational and performance reporting needs. Reliable data is vital to the processing and servicing of loans and grants, and the oversight of campus-based, Pell, Direct Loan and FFEL programs. High quality data also provides insight into operations and improves the decision making process. The following policies implement a collaborative process between business owners and the EDM Team for assessing, monitoring and improving the quality of mission-critical business data:

- **Conduct routine data quality monitoring**: Business owners will submit to the EDM Team annually or upon request electronic data copies from their systems. The EDM Team will then perform data profiling and will generate detailed reports for the business owners. The profiling and reporting will be based on a set of metrics tailored for specific business owner needs. To facilitate the implementation of this policy, the EDM Team will maintain a data quality toolset and offer it as a service to the enterprise. Refer to Section 3 for a list of EDM services.

- **Analyze data quality reports and take corrective actions if necessary**: After data profiling is completed, the EDM Team will report data quality findings back to the business owners, who then will decide if they need to take action based on their own established data quality standards and policies, if any. Business owners can also utilize EDM-provided data quality best practices to examine their existing data quality processes, policies and practice, or to establish new ones. To facilitate the implementation of this policy, the EDM Team will develop up to date knowledge and expertise in data quality best practices and offer consultancy services to the enterprise. Refer to Section 3 for a list of EDM services.

4.3 Data Governance Program

The management of data across the enterprise relies on commonly agreed-upon data definitions. Data governance defines processes and procedures for reaching these agreements. Currently, the EDM Team is defining an enterprise data governance plan based on industry best practices. The following are future policies after the data governance program is established.

- **Participate in the enterprise data governance program**: Business owners will participate in the enterprise data governance program and will represent relevant BCAs in the decision making process.

- **Assign enterprise data stewardship**: Business owners will designate data stewards from their BCAs. The data stewards will have day-to-day responsibility for coordinating data governance activities.

4.4 Use of XML R&R for the Education Community

The Postsecondary Education Standards Committee (PESC) XML R&R was created for the Education Community. It was developed and is operated by the US Department of Education's Office of Federal Student Aid and is administered by PESC and Federal Student Aid. The XML R&R is a central access point for XML core components, XML schemas and supporting documentation for the education standards community, of which Federal Student Aid is member. The repository was created and reviewed through a collaborative effort between Federal Student Aid, PESC and the education standards community. Therefore, its contents reflect agreement reached among a diversified community and are widely accepted. The XML R&R can be accessed at www.FSAxmlRegistry.ED.gov. The following policies govern the use of the XML R&R at Federal Student Aid:

- **Use XML R&R for the Education Community as the enterprise XML registry and repository**: Until an enterprise MDR solution is put in place at Federal Student Aid (refer to Subsection 4.6 for more detail), the PESC XML R&R will serve as the enterprise XML registry and repository to support all Federal Student Aid XML development needs. Adopting PESC XML R&R ensures all Federal Student Aid XML development activities follow the same standards, and it sets a foundation for system interoperability among internal applications and between internal and external applications. Federal Student Aid project managers and data architects will use the PESC XML R&R as the only source for common XML data elements to be implemented in their development initiatives. Data architects will use the PESC XML R&R to search and retrieve the latest definitions of standardized XML elements.

- **Follow the established policies and procedures for XML registry and repository changes**: Project managers and data architects also may request additions and modifications when the XML R&R for the Education Community does not meet their needs. Change requests must follow the established PESC policies and procedures, which can be accessed at http://www.pesc.org/info/policies-procedures.asp[4].

4.5 Use of the ECDM, ELDM and Enterprise Naming Standards

The ECDM captures high-level data entities and their relationships. The ECDM is used to depict general boundaries between BCAs from a data perspective. The ELDM further defines the data entities and relationships. Both ECDM and ELDM are developed in business terms and are used to capture data requirements and business rules, and to publish common data definitions used at Federal Student Aid. These models are the foundation for data design activities for any development project. The following policies promote the use of the ECDM and ELDM in all development initiatives as their usage facilitates implementation of common data architecture:

[4] Postsecondary Electronics Standards Council (PESC) Policy and Procedures Manual, http://www.pesc.org/info/policies/Manual-Finalv1-0-0.doc, April 2005.

- **Use the ECDM and ELDM in all data design activities**: Data architects will directly adopt the ECDM and ELDM for capturing data requirements and creating project data models. The EDM Team will maintain the ECDM and ELDM, provide them upon request to development teams, and assist data architects in understanding and adopting them.

- **Follow Enterprise Operation Change Management (EOCM) procedures for ECDM and ELDM changes**: Data architects and DBAs will submit change requests to the EOCM board if they find in their analysis and design that the models do not meet business user requirements. After the EOCM evaluates and approves a proposed change, the EDM Team will update the models accordingly.

- **Use ELDM as the basis for Inter-system data exchanges**: Requests made by one system to another for certain data elements should be made with reference to data entities and attributes on the ELDM. This practice ensures enterprise interoperability and data consistency.

- **Adopt the enterprise object naming standards**: The EDM Team is developing enterprise data object naming standards. Once the standards are finalized, data architects and DBAs will follow those standards in naming logical and physical data elements.

4.6 Enterprise Metadata Repository

An enterprise MDR collects and publishes standard data definitions, links data definitions to data element locations and captures data element movements. The MDR provides an up-to-date snapshot of Federal Student Aid data assets to facilitate data analysis and design. An MDR requires implementation of a metadata solution through custom development or through purchase of a software package. The following policies apply only after Federal Student Aid has such a metadata solution in place:

- **Use the enterprise MDR as the project metadata source**: Data architects and DBAs will use the enterprise MDR for data analysis and for design activities by retrieving and using such artifacts as data dictionary and XML schemas. When these artifacts do not meet their specific requirements, data architects and DBAs will submit change requests to the EOCM for operational issues, and to Enterprise Governance, if developed, for requirements related issues.

- **Keep the MDR updated with latest metadata information**: Data architects and DBAs also will provide to the EDM Team quarterly updates to various metadata in the MDR. The following are examples of such metadata:

    a. Data sources
    b. Data movements and mappings
    c. Transformation
    d. Usage of a data element by business processes, systems, operational reports and/or BI analyses.

4.7 Implementation of Enterprise MDM

Federal Student Aid defined a long-term architecture vision called the Enterprise Vision (EV). It is essential that the EV standards are followed by development projects for Federal Student Aid to move towards integrated business processes and systems architecture. A set of service-oriented architecture MDM components complements the TSV business and systems architecture with integrated enterprise data architecture. The following policies apply to the implementation of these MDM components to maintain data integrity and to ensure data consistency across the enterprise:

- **Implement enterprise MDM services to manage systems of record for key data elements, such as Person, Organization and Aid**: Various systems will store the master records for specified sets of common data elements (e.g., person, organization, aid). Data architects will ensure that their solutions contain all necessary components to manage the integrity of such common data elements. For the master data assigned to a system, the solution will provide MDM services to share the data with other systems and to manage update requests from other systems. For specific types of MDM services and their functions, please refer to relevant EV documentation.[5]

- **Conform to enterprise standards in designing data interfaces**: Data architects will ensure that data interface design adheres to the technical standards specified by the EV integration strategy. For example, intra-system communications and data exchanges are through common services accessible on the Enterprise Service Bus (ESB). Please refer to relevant EV documentation for more detail.

- **Adopt COTS products that conform to Federal Student Aid data integration standards**: COTS products will adhere fully to Federal Student Aid EV integration strategy and MDM standards. Project managers and data architects will evaluate all COTS products to ensure that:
    - They can integrate with the EV for all intersystem communications;
    - They can provide an MDM service interface that exposes data from its proprietary data store(s) and that allows creation, updates and deletes of data elements;
    - The COTS proprietary data can be mapped to standard Federal Student Aid ELDM data entities.

4.8 Data Security and Privacy

Protection of data and customer privacy is a top priority of Federal Student Aid. Any system initiatives must comply with Federal laws, regulations and policies, and with Department of Education security policies. In addition, Federal Student Aid is currently refining security and privacy standards. While this effort is going on, the following policies govern the data design for security and privacy requirements:

- **Manage data as a key information resource**: The Department of Education Handbook of Information Assurance Security Policy[6] sets forth policies in protecting the confidentiality, integrity, and availability of the Department's information. Business owners and project managers will treat data as a key information resource, and manage data according to those policies established in the Handbook.

- **Follow all relevant laws, regulations and standards**: Data architects will ensure their solutions follow relevant laws and regulations, and Federal Student Aid standards, regarding the design and management of access controls, enterprise roles and audit logging. Refer to Target State Vision Security Policies[7] and Target State Vision Security Standards for a list of these laws, regulations and standards.[8]

- **Address security and privacy protection in data design**: Data architects will ensure that data design address the following areas:

[5] Information Framework, December 2005.

[6] Department of Education Handbook of Information Assurance Security Policy, March 2007.

[7] Target State Vision Security Policies Draft, November 2005.

[8] Target State Vision Security Standards Draft, November 2005.

- Data risk analysis and data classification, which defines the security and privacy requirements for each data element.
- Data view administration, which defines and manages logical representations of the raw, physical data and thus limit its exposure.
- Data access permission administration, which manages access levels for different users based on their roles and business needs.
- Data in transmission security, which focuses on security and privacy issues associated with data while it is being transferred, physically or electronically.
- Audit capabilities, which enable the creation, retention, management and usage of audit trails.

- **Follow enterprise security architecture standards and guidelines**: Data architects will ensure data solutions follow the architecture and standards specified in the Federal Student Aid Identify and Access Management Vision[9].

[9] Identify and Access Management Vision, May 2007.

Appendix A. Acronyms

The following acronyms are used in this document.

Acronym	Definition
BCA	Business Capability Area
BI	Business Intelligence
CCA	Clinger-Cohen Act
CONOP	Concept of Operations
COTS	Commercial Off The Shelf
DBA	Database Administrator
DRM	Data Reference Model
ECDM	Enterprise Conceptual Data Model
ED	Department of Education
EDW	Enterprise Data Warehouse
ELDM	Enterprise Logical Data Model
EOCM	Enterprise Operation Change Management
ESB	Enterprise Service Bus
EV	Enterprise Vision
FEA	Federal Enterprise Architecture
FFEL	Federal Family Education Loans
IT	Information Technology
MDM	Master Data Management
MDR	Metadata Repository
NSLDS	National Student Loan Data System
ODS	Operational Data Store
OMB	Office of Management and Budget

Acronym	Definition
PESC	Postsecondary Electronic Standards Council
R&R	Registry and Repository
SDLC	System Development Life Cycle
SOA	Service Oriented Architecture
TSV	Target State Vision
XML	eXtensible Markup Language

Table A-1. Acronyms